NOT TO BE FURTHER COPIED OR REPRODUCED

(* y VS. x *)

14

Bends Are Also There To Be Taken Away

		Exh.	Page
Introduction	14-1		
		A	
		1	14-3
		2	14-4
		3	14-6
14A. The vapor pressure of water	14-2		
		B	
14B. Assessing bend from three points	14-5		
- arithmetic of fitting	14-8		
		C	
		4	14-13/14
		5	14-15/16
		6	14-18
		7	14-19
		8	14-21
14C. A simple bend assessor	14-10		
- braking again	14-11		
14D. Mopping up small bends	14-20		
- marriages and m. v. deaths again	14-20		
		D	
		9	14-25
- the functional example again	14-23	10	14-26
		11	14-28
* * *			
		H	
		K	
		P	
14H. How far have we come?	14-29		
* * *			
14K. Some touches of algebra (optional background)	14-3		
- why quadratics?	14-31		
- difference in tilts for x^2	14-32		
* * *			
14P. Problems	14-34		

14

Bends Are Also There To Be Taken Away

Now that we have learned to deal fairly effectively with tilt, it is time to go on and deal with something further. Just as we discussed the fitting and subtraction of straight lines in terms of tilt and level, so now we will discuss the fitting and subtraction of simple curves in terms of bend, tilt, and level.

Just as we tried to settle with tilt first, leaving level to be settled later, so too will we try to start with bend, then turn to tilt, and finish with level. For tilt, such a simple program worked cleanly and effectively, mainly because fitting for level influences none of the calculations we might want to apply in other than the simplest way. This is not always the case when we are dealing with bend -- subtracting a new trial tilt can, for example, modify middle means of portions in other than a simple way. We will do what we can, however, to come close to the simple rule of "fitting the most complex thing first" .

The only approach to arithmetic fitting of lines that we discussed was the two-step approach of:

- find representative or central points,
- pass fit through points,

though other methods of fitting lines are frequently used. We will follow the same sort of approach about bend, using three representative or central points rather than two.

14A. The vapor pressure of water

As a first example, let us turn back to the vapor pressure of water (see section 9E) and work in terms of selected points on graphs, leaving all details of calculation aside. Our first purpose is to show what can be done, not how to do it.

Exhibit 1 repeats exhibit 16 of chapter 9, showing an untilted form of the log vapor pressure of water, with the addition of:

- three selected points.
- a standard bend through these points.

Exhibit 2 shows the further residuals obtained by subtracting this standard bend from the points. We now see considerably more than we did before. The break at 0°C -- corresponding to the difference between liquid water and solid water -- is now much more prominent. Moreover, we see that our bend, chosen to fit more or less overall, does not match the bend for liquid water very closely.

To see more, we need to fit a bend to the liquid water portion of the data. It would be nice if we could do this by merely fitting another standard bend on top of the one we have already fitted. Gaining the ability to do this would be a good reason for choosing a particular family of standard bends.

In fact, the standard bends we are using do have this property. Choosing three points on the liquid water curve and subtracting the standard bend through them gives

exhibit 1 of chapter 14

Residuals of log vapor pressure of water from a straight line
(Repeat of exhibit 16 of chapter 9.)

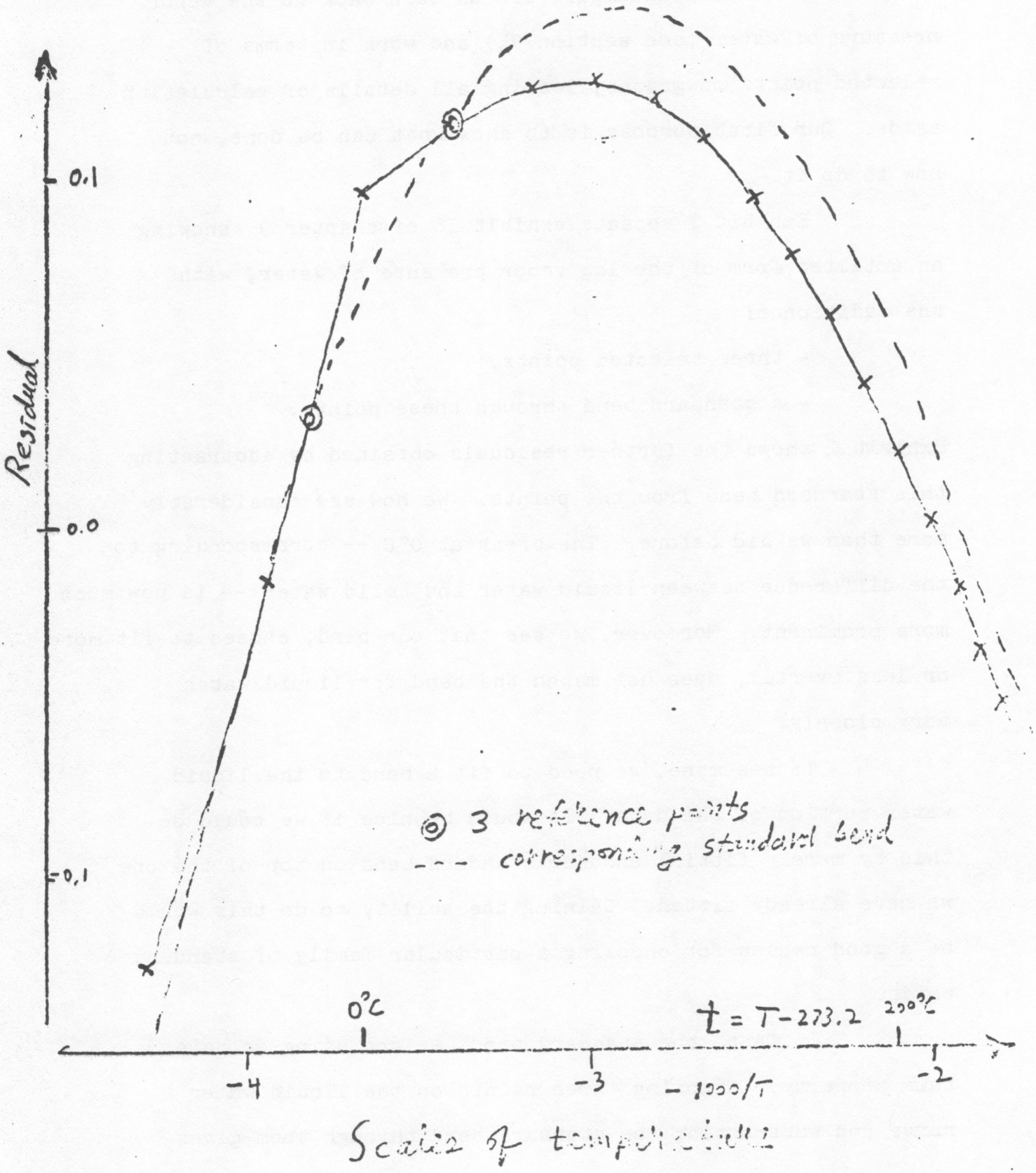


exhibit 2 of chapter 14

Residuals of log vapor pressure of water after fitting a quadratic in -1/T to three wide-spread points

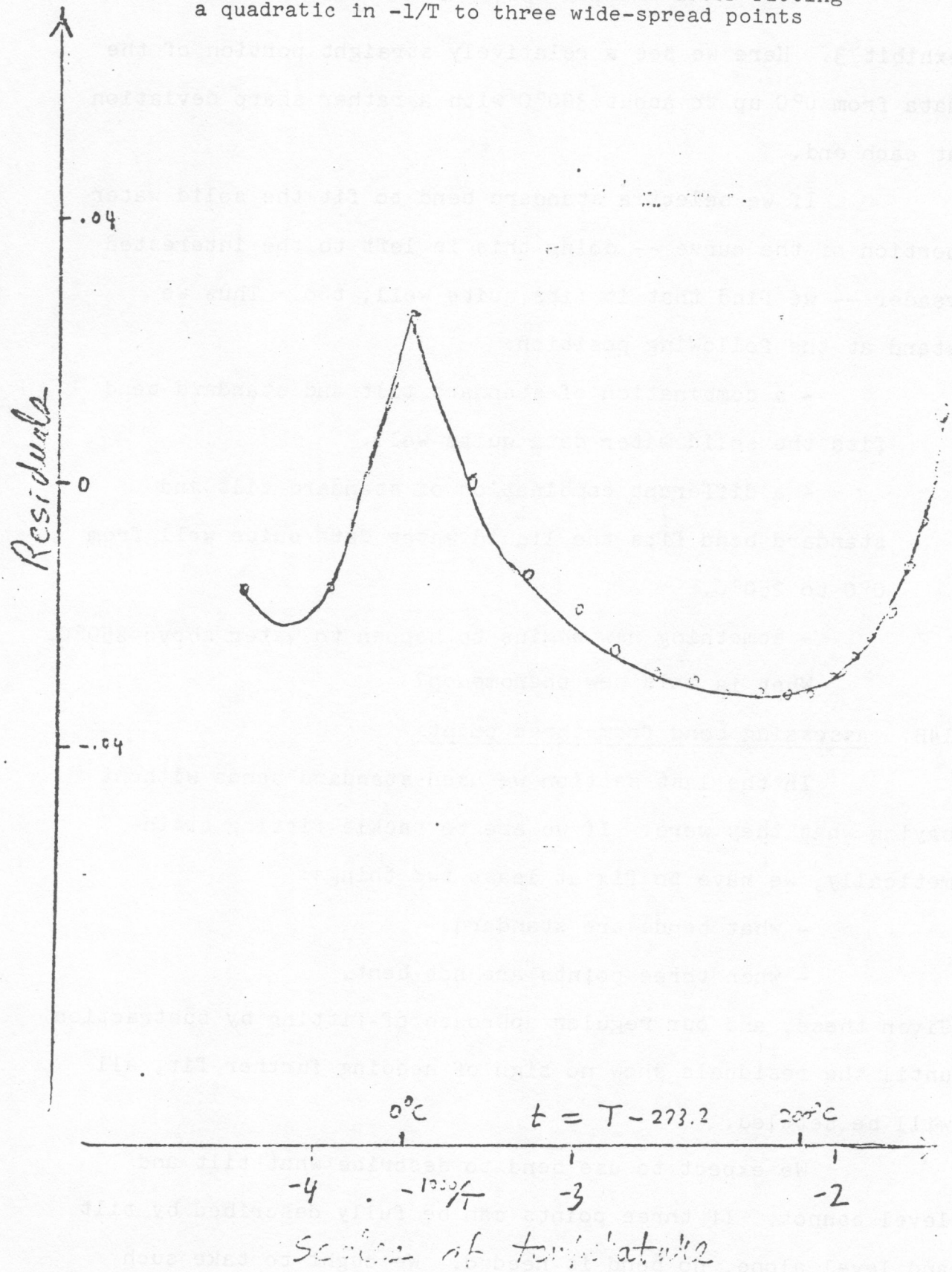


exhibit 3. Here we see a relatively straight portion of the data from 0°C up to about 300°C with a rather sharp deviation at each end.

If we select a standard bend to fit the solid water portion of the curve -- doing this is left to the interested reader -- we find that it fits quite well, too. Thus we stand at the following position:

- a combination of standard tilt and standard bend fits the solid water data quite well.
- a different combination of standard tilt and standard bend fits the liquid water data quite well from 0°C to 250°C.
- something new begins to happen to water above 250°C. What is this new phenomenon?

14B. Assessing bend from three points

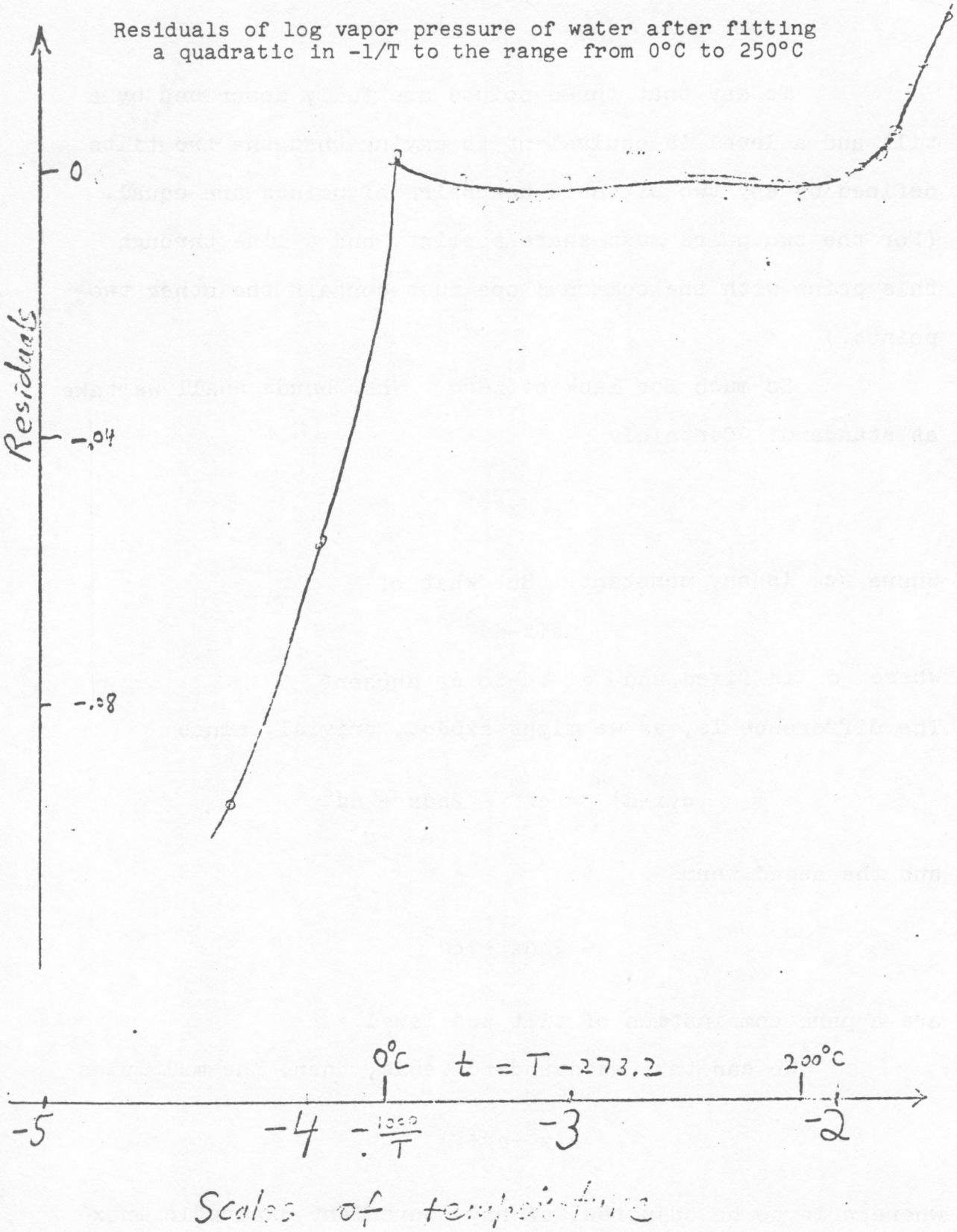
In the last section we used standard bends without saying what they were. If we are to tackle fitting arithmetically, we have to fix at least two things:

- what bends are standard.
- when three points are not bent.

Given these, and our regular approach of fitting by subtraction until the residuals show no sign of needing further fit, all will be settled.

We expect to use bend to describe what tilt and level cannot. If three points can be fully described by tilt and level alone, no bend is needed. We ought to take such situations as defining "no bend" for three points.

Residuals of log vapor pressure of water after fitting a quadratic in -1/T to the range from 0°C to 250°C



To say that three points are fully described by a tilt and a level is equivalent to saying that the two tilts defined by any two of the three pairs of points are equal. (For the two pairs must share a point, and a line through this point with the common slope must contain the other two points.)

So much for lack of bend. What bends shall we take as standard? Certainly

$$cx^2$$

where c is any constant. But what of

$$c(x-d)^2$$

where d is fixed and c is to be chosen?
The difference is, as we might expect, trivial, since

$$c(x-d)^2 = cx^2 - 2cdx + cd^2$$

and the added terms

$$- 2cdx + cd^2$$

are a pure combination of tilt and level.

We can take as standard bends, then, the multiples

$$c(x^2+ex+f)$$

where c is to be adjusted, of any convenient quadratic in x

$$x^2+ex+f$$

where e and f have any fixed values.

* arithmetic of fitting *

Let us turn then to the practice of fitting. The easiest case to treat, of course, is one where three points are given. Suppose, then, that we are given these three points:

x	y
3	7
5	4
8	11

How are we to assess the amount of bend corresponding to these points? We surely wish that our assessment of bend should avoid any response to pure tilt. What is an easy way to do this?

One easy approach is to use the three points to calculate two different tilts, and then depend on the

difference between these tilts

to assess the bend. For our example, we have -- for our data and for x^2, a standard bend:

x	y	tilt
3	7	
		1.50
5	4	
		1.33
8	8	
diff		2.83

x	x^2	tilt
3	9	
		8.00
5	25	
		13.00
8	64	
		5.00

Since the observed difference in tilt is

$$\frac{2.83}{5.00} = .567$$

times as large for the data as for x^2, the natural assessment of bend corresponds to a bend described by $.567x^2$. Trying this out we find:

x	y	$.567x^2$	$y-.567x^2$	tilt
3	7	5.10	1.90	
				-6.04
5	4	14.17	-10.17	
				-6.03
8	8	36.27	-28.27	
	(data)	(bend)	(residual)	

Clearly this process has provided three residuals related by pure tilt and level.

In doing this example, we used x^2 as an example of a standard bend. There was no need for making this specific choice. If, as another example, we had chosen $(x-5)^2$, we would have found:

x	y	tilt	x	$(x-5)^2$	tilt
3	7		3	4	
		-1.50			-2.0
5	4		5	0	
		+1.33			+3.0
8	8		8	9	
diff:		2.83			5.0

again finding 0.567 as the coefficient, here of $(x-5)^2$. Continuing, we would find

x	y	$0.567(x-5)^2$	$y - 0.567(x-5)^2$	tilt
3	7	2.27	4.73	
				-0.37
5	4	0	4.00	
				-0.37
8	8	5.10	2.90	

again finding three residuals which represent pure tilt and level. (We expect this tilt and level to be different from that found by subtracting $0.567x^2$.)

Clearly either approach leads to a satisfactory answer and is equally useful. Equally, we could have used $(x-7)^2$, or $(x-\pi)^2$, etc. The only reason for one choice over another is simplicity of arithmetic. From this point of view, a choice of $(x-d)^2$ with d somewhere interspersed among the values of x leads to smaller numbers and thus, usually, simpler arithmetic. In our example, the choice $d = 5$ obviously works well.

Sometimes -- as we will soon see in the case when d is a half-integer, but all the x's are integers -- it may simplify the arithmetic a little to use $(x-d)^2 - e$ rather than $(x-d)^2$. We are clearly free to do this, since doing this will affect the calculated slopes not at all, and will lead to the same amount of bend.

14C. A simple bend assessor

If we have more than three points, as we usually do, we need to begin with some form of summarization of the data. Again we can follow the broad general approach we used for tilt.

This time we break our observations up into five portions according to their x-values, and calculate the apparent direction and amount of bend from the behavior of central points for three of these: the lowest, marked L; the middle, marked M; and the highest, marked H. (When we come to set a level, we can use central points for the other two as well.

A convenient rule of thumb about which points belong to these three portions runs as follows:

- end portions begin at an extreme and extend less than 1/2 the way to the median.
- the middle portion surrounds the median and extends less than 1/2 the way to either extreme.
- all points with a given x either belong to a given portion or none do.
- no one of the three portions contains more than 1/5 the points.
- subject to all these constraints, each of the three portions contains as many points as it can.
- if there remains a choice for the center portion, that choice is taken which centers the points best

a around the median.

* braking again *

For an example, we can turn back to the braking example of section 9D. Exhibit 5 of chapter 9 gives 50 combinations of speed and distance to stop. The speeds run from 4 to 25 mph, with a median of 15.

The first two parts of the rules of thumb now tell us that the three main portions are to be confined to the following intervals of x:

4 to 9	(6 points available)
10 to 19	(32 points available)
21 to 25	(7 points available)

Since no more than 50/5 = 10 points are available for the two end portions, we take them all. For the middle portion we must cut 32 down to 10. If we take x = 14, 15, and 16, we get

$$4+3+2 = 9 \text{ points}$$

Clearly this is as good centering as we can get.

Exhibit 4 shows the data separated into portions and the analysis begun. We naturally try a bend of $0.1x^2$, in view of the apparent bend of $0.09x^2$.

Exhibit 5 proceeds to fit a better bend -- and look further -- then a tilt, and then a polish to both bend and tilt before calculating a level. A careful worker would -- and we will -- calculate summary residuals for the remaining two portions before stopping.

We have come far enough, however, to make one interesting and important point. We have just fitted

$$y \sim .08x^2 + 1.58x - 3.5$$

exhibit 4 of chapter 14

The Braking Data Examined for Bend, 1

(see exhibit 5 of chapter 9)

A) DATA in PORTIONS and FIRST SUMMARY POINTS

x Speed (mph)	y Distance (feet)	Trimeans for x	Trimeans for y
4	2,10,		
7	4,22,		
8	16,	6.5	10.0
9	10,		
10	18,26,34,		
11	17,28,		
12	14,20,24,28,	12.0	26.5
13	26,34,34,46,		
14	26,36,60,80		
15	20,26,54,	14.8	38.0
16	32,40,		
17	32,40,50		
18	42,56,76,84,		
19	36,46,68,	19.0	51.0
20	32,48,52,56,64,		
22	66,		
23	54,		
24	70,92,93,120,	23.8	82.2
25	85,		

B) CALCULATION OF FIRST BEND

x	y	tilt	diff	x^2	tilt	diff	ratio
6.5	10.0			42.25			
14.8	38.0	3.37	1.54	219.04	21.3	17.3	0.09
23.8	82.2	4.91		566.44	38.6		

Apparent bend = $0.09x^2$

Trial bend = $0.1x^2$

(Cont'd)

exhibit 4 of chapter 14 (Cont'd)

Examples: $3.37 = \frac{38.0-10.0}{14.8-6.5}$, $4.91 = \frac{82.2-38.0}{23.8-14.8}$

$21.3 = \frac{219.04-42.25}{14.8-6.5}$, $38.6 = \frac{566.44-219.04}{23.8-14.8}$

$0.09 = \frac{1.54}{17.3}$

C) LABOR-SAVING NOTE

Notice that also,

$$21.5 = 6.5+15.0, \quad 38.8 = 15.0+23.8$$

so that

$$17.3 = 38.8-21.5 = 23.8+15.0-15.0-6.5$$
$$= 23.8-6.5$$

The general formula is:

$$\text{difference in tilts of } x^2+dx+e = x_H-x_L$$

Its use can save us considerable arithmetic.

exhibit 5 of chapter 14

The Braking Data Examined for Bend, 2

A) INITIAL CALCULATIONS

x	$.1x^2$	$y-.1x^2$	$-.03x^2$	(*)	Diffs	2x	Diffs
4	1.6	0,8,	-0.5	-1	1,9,	8	-7,1,
7	4.9	-1,17,	-1.5	-2	1,19,	14	-13,5,
8	6.4	10,	-1.9	-2	12,	16	-4,
9	8.1	2,	-2.4	-2	4,	18	-14,
14	19.6	6,16,40,60,	-5.9	-6	12,22,46,66,	28	-16,-6,18,38,
15	22.5	-2,4,32,	-6.8	-6	4,10,38,	30	-26,-20,8,
16	25.6	6,14,	-7.8	-8	14,22	32	-18,-10,
22	48.4	18,	-14.5	-14	32,	44	-12,
23	52.9	1,	-15.9	-16	17,	46	-29,
24	57.6	12,34,35,62,	-17.3	-18	30,52,53,80	48	-18,4,5,32
25	2.5	23,	-18.8	-18	44,	50	-9,
trimeans		5.0			6.5		-5.8
		16.5			23.5		-7.5
		23.2			41.2		-7.8

See below for explanation of (*), which is clearly close to $-.03x^2$.
Diffs are $y-.1x^2$ minus (*) .

B) CALCULATION of 2ND TRIAL BEND

x	$y-.1x^2$	tilt	diff	ratio
6.5	5.0			
		1.39		
14.8	16.5		-.65	-.038
		.74		
23.8	23.2			

Ratio is to 17.3, as before.
2nd trial bend = -.03.

C) CALCULATION of 3RD TRIAL BEND

x	$y-.07x^2$	tilt	diff	ratio
6.5	6.5			
		2.05		
14.8	23.5		-.08	-.005 (take it)
		1.97		
23.8	41.2			

Trial tilt after 2nd bend = 2.0.

(Cont'd)

exhibit 5 of chapter 14 (Cont'd)

D) CHECK AFTER TRIAL TILT

x	$y-.07x^2-2x$	tilt	diff	ratio	$.01x^2$	$y-.08x^2-2x$	tilt
6.5	-5.8				0.4	-6.2	
14.8	-7.5	-0.20	.17	.01	2.2	-9.7	-.42
23.8	-7.8	-0.03			5.7	-13.5	-.42

Polish = $.01x^2 - .42x$

E) CALCULATION of LEVEL

x	-.42x	$y-.08x^2-1.58x$	-3.5	Residual
6.5	-2.7	-3.5	-3.5	0.0
14.8	-6.2	-3.5	-3.5	0.0
23.8	-10.0	-3.5	-3.5	0.0

Fit: $y \sim .08x^2 + 1.58x - 3.5$

F) EXPLANATION of (*)

The entries in column (*) are the integers that come closest to adjusting $y-.1x^2$ to $y-.07x^2$, when both are taken to the nearest integer. Thus, for example,

$.1x^2$	applied	$.07x^2$	best applied	same - applied = (*)
1.6	2	1.1 = 1.6-0.5	1	-1
4.9	5	3.4 = 4.9-1.5	3	-2
6.4	6	4.5 = 6.4-1.9	4	-2
8.1	8	5.7 = 8.1-2.4	6	-2
19.6	20	13.7 = 19.6-5.9	14	-6
22.5	22	15.7 = 22.5-6.8	16	-6
25.6	26	17.8 = 25.6-7.8	18	-8

By the close of section 9E we had fitted

$$y \sim .115x^2 + 0.7x$$

Which fit should we prefer?

The controlling fact is our knowledge that at 0 miles per hour it takes exactly 0 feet to stop, every time. That fact has not been used in the present section. As a result, we have fitted a quadratic that says we will stop in -3.5 feet at 0 miles per hour. What brakes! (Do they use jets -- or rockets?)

Had we not known this simple fact, the fit of this section would have been a reasonable fit -- when reasonable is interpreted in its proper meaning -- "we don't know how to do much better" . As it is, we have to prefer the fit of section 9E.

Exhibit 6 goes back and checks the two remaining portions. One fits nicely, the other is 5 feet short. If we had not recalled our anchor point at zero, this low residual might have caused us to shift the level of the fit by -0.9.

Exhibit 7 shows the anchor point and the five summary points, compared to both fits. The result of this section clearly fits closer over most of the interval of speeds. Whether or not the apparent lower precision of fit of the formula of section 9E should concern us is a matter of confirmatory data analysis, and thus formally outside our scope here. We do have a righ

exhibit 6 of chapter 14

The Braking Data, 3

Checking the Other Two Portions

A) DATA and RESIDUALS

x	$.08x^2$	1.58x	-3.5	$.08x^2+1.58x-3.5$	Residuals
10	8.0	15.8	-3.5	20.8 → 21	-3,5,13,
11	9.7	17.4	-3.5	23.6 → 24	-7,4,
12	11.8	19.0	-3.5	27.3 → 27	-13,-7,-3,1,
13	13.5	20.5	-3.5	30.5 → 30	-4,4,4,16,
17	23.1	26.9	-3.5	46.5 → 46	-14,-6,4,
18	25.9	28.4	-3.5	50.8 → 51	-9,-5,25,33,
19	28.9	30.0	-3.5	55.4 → 55	-19,-9,13,
20	32.0	31.6	-3.5	60.1 → 60	-28,-12,-8,-4,4,
				trimeans	0.5
					-5.0

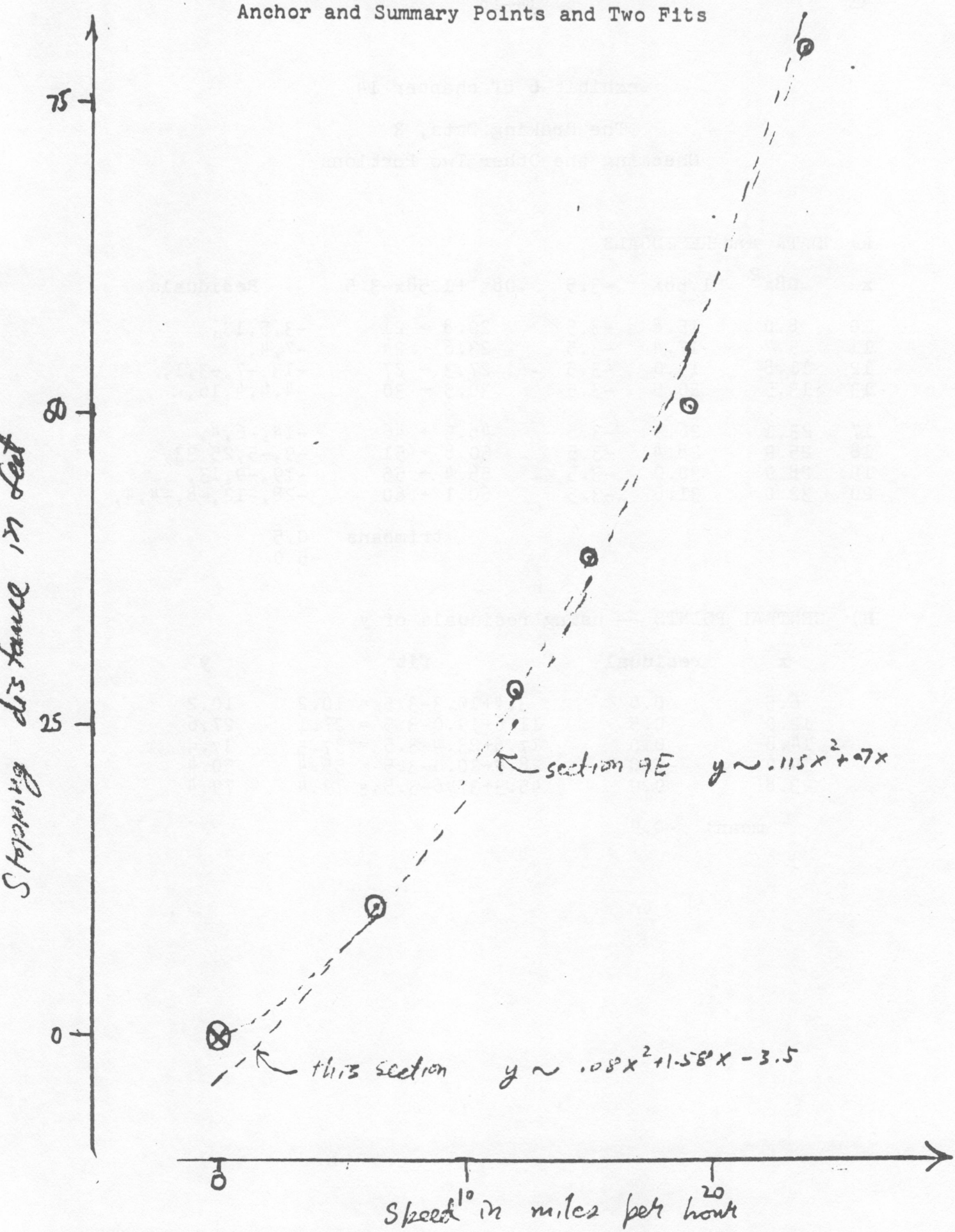
B) CENTRAL POINTS -- using residuals of y

x	residual	fit	y
6.5	0.0	3.4+10.3-3.5 = 10.2	10.2
12.0	0.5	11.6+19.0-3.5 = 27.1	27.6
14.8	0.0	17.6+23.4-3.5 = 37.5	37.5
19.0	-5.0	28.9+30.0-3.5 = 55.4	50.4
23.8	0.0	45.3+37.6-3.5 = 79.4	79.4

mean: -0.9

exhibit 7 of chapter 14

Anchor and Summary Points and Two Fits



to look and see what we can, however, so exhibit 8 shows -- in residual form -- the observations and the two fitted quadratics. It is reasonably clear that -- except for the firmly known anchor point at (0,0) -- the difference between the two fits is surprisingly small compared to the scatter of the observations.

There seems to be no reason to avoid the earlier fit -- the one that made use of our knowledge of (0,0).

14D. Mopping up small bends

Rules of thumb are to help us, not to constrain us. If we are offered three summary points on a silver platter, we should use them first and ask questions afterward.

* marriages and m.v. deaths again *

Let us turn back to exhibit 12 of chapter 10, relating

$$y \sim \text{log of motor vehicle deaths}$$

to

$$x \sim \text{log of marriages}$$

for 49 states (including the District of Colombia).

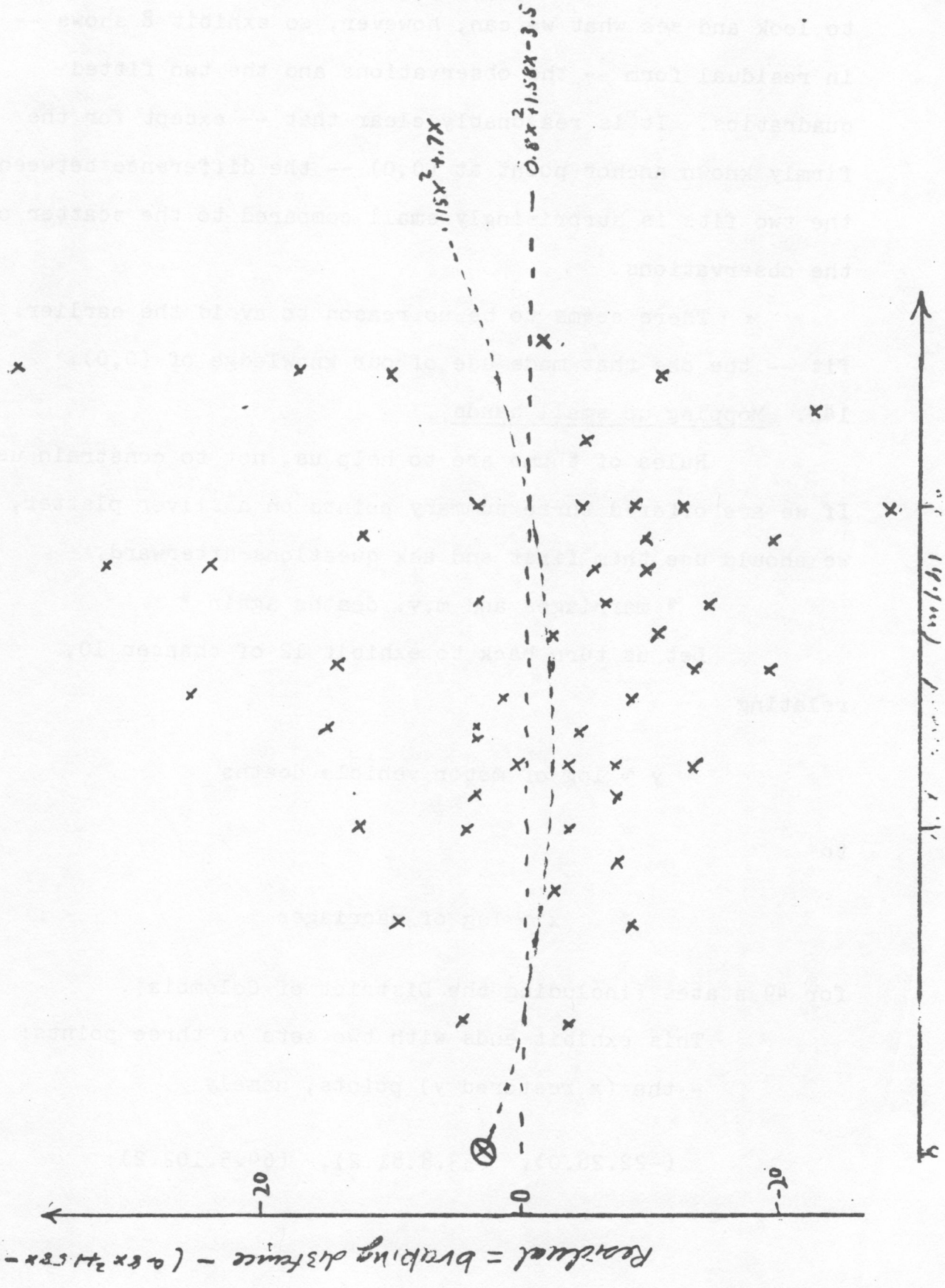
This exhibit ends with two sets of three points:

- the (x,restored y) points, namely

(-22,28.0), (33.8,81.2), (69.5,102.2)

exhibit 8 of chapter 14

The Residuals of the Data and the Residuals of the Two Fitted Quadratics



- the (x,polished residual) points, namely

$$(-22,-2.66),\quad (33.8,5.33),\quad (69.5,-2.66)$$

Since these two sets differ only by a fitted straight line, we would be foolish -- if doing hand arithmetic -- to work with the larger numbers of the first set.

The second set gives tilts of

$$\frac{5.33 - (-2.66)}{33.8 - (-22.\ )} = \frac{8.0}{55.8} = .14$$

and

$$\frac{-2.66-5.33}{69.5-33.8} = \frac{-8.0}{35.7} = -.22$$

whose difference $-.36$ is to be compared with the difference of tilts for x^2, here given by $69.5 - (-22.\ ) = 91.5$. Thus these points would be unbent by subtracting

$$-.0049x^2$$

and, as the reader may verify, can be fitted -- to reasonable precision -- by

$$-.004x^2 + .19x + 3.5$$

We have done this fit using summary points only. How much trouble could this have brought us? The extreme ends of the lower portion were -63 and 0, for which $-.004\ x^2$ takes the values -13.8 and 0, while $-.004\ x^2 + .19x$ goes further, to -27.8 and 0. If we were to take this quadratic

term seriously -- which we shall not do -- its effect on the individual points is large enough so we cannot be sure that we dare neglect it.

* the functional example again *

Let us turn next to exhibit 1 of chapter 10, which began our adventures in trying to get six-place answers by fitting a five-place table of a certain function. The three (x,residual) points we found there were:

(.006,.000003), (.025,-.000007), (.044,.000003)

for which the slopes are

$$\frac{-.000010}{.019} = -.00053$$

and

$$\frac{+.000010}{.019} = .00053$$

with a difference of .00105. The difference for x^2 is

$$.044 - .006 = .038$$

so the indicated bend is

$$.028x^2$$

or, as is likely to be more convenient

$$.028(x-.025)^2$$

Calculating this latter bend to seven decimals and subtracting from the residuals of exhibit 17 of chapter 10

leads to the results set out in exhibits 9 (numbers) and 10 (plot). The eye-fit can now be made much more precisely, especially when we use "windows" for the admitted and inevitable inaccuracies of up to ±.000005 that must exist in five-decimal values.

The windows in exhibit 10 mark out places where a fit would leave a final residual of more than ±.000005. If all were well, we ought to be able to draw in a fitting line which avoids all of these barriers. Line A does this.

However, line B -- expressable as

$$-.000004 - .0001x$$

-- seems to fit the points better, although crossing a single barrier. Using this line corresponds to using the sum of

$$.05139x + .398929$$

$$.028(x-.025)^2$$

$$-.000004 - .0001x$$

which is

$$.398925 + .05129x + .028(x-.025)^2$$

or, alternatively,

$$.398942 + .04989x + .028x^2$$

as our fit to the fifteen 5-decimal values of the function.

The proof of the pudding, however, is here to be found by comparing our final six-decimal fits with the true six-decimal values of the function.

exhibit 9 of chapter 14

Further Adjustment of Residuals in the Functional Example

(see exhibit 17 of chapter 10)

A) CALCULATIONS

x	old residual	$.028(x-.025)^2$	Diff	line* new fit*	10^6x Diff
.000	.000011	.000018	-.000007	-.000004	-3
.004	.000005	.000012	-.000007	-.000004	-3
.006	.000002	.000010	-.000008	-.000005	-3
.008	.000000	.000008	-.000008	-.000005	-3
.014	.000001	.000003	-.000002	-.000005	3
.016	-.000001	.000002	-.000003	-.000006	3
.020	-.000007	.000001	-.000008	-.000006	-2
.026	-.000005	.000000	-.000005	-.000007	2
.028	-.000008	.000000	-.000008	-.000007	-1
.030	-.000011	.000001	-.000012	-.000007	-4
.040	-.000005	.000006	-.000011	-.000008	-2
.042	.000002	.000008	-.000006	-.000008	2
.044	.000000	.000010	-.000010	-.000008	-2
.046	.000007	.000012	-.000005	-.000009	4
.050	.000012	.000018	-.000006	-.000009	3

* Obtained graphically in exhibit 10.

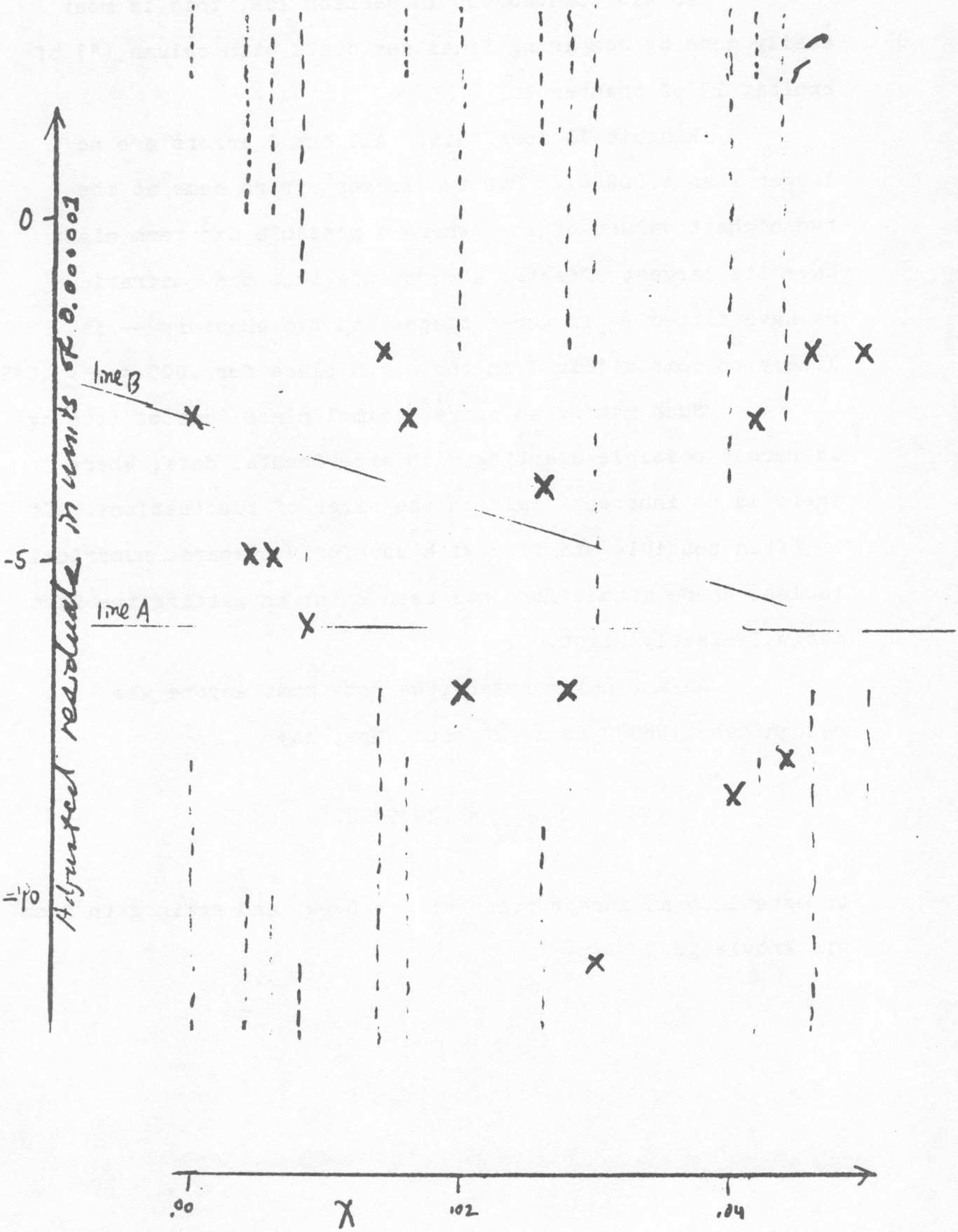
B) SUMMARY OF FINAL DIFF'S

-5	5
-4	
-3	333
-2	222
-1	11
0	
1	
2	22
3	33
4	44
5	

exhibit 10 of chapter 14

Adjusted Residuals and Their Windows

(see exhibit 9)



As was pointed out in section 10K, this is most easily done by comparing final residuals with column (*) of exhibit 19 of chapter 10.

Exhibit 11 does this. All but 2 errors are no larger than ±.000001. The two larger errors come at the two highest values of x -- where a possible dx^3 term might have its largest effect. We conclude that the quadratic we have fitted -- in three stages and two chapters -- is likely to come within 1 in the sixth place for $.000 \leq x \leq .045$.

Such making an extra decimal place "out of nothing" is rarely possible starting with experimental data, where there is no inherent limit on the sizes of fluctuations. It is often possible starting with carefully prepared numerical tables, where much effort has been spent on getting the last decimal exactly right.

As a closing remark, we note that anyone who recognizes .398942 as $1/\sqrt{2\pi}$ could use, say

$$\frac{1}{\sqrt{2\pi}} = .39894\ 228$$

to establish an anchor point at x = 0.000 and again gain from his knowledge.

exhibit 11 of chapter 14

Did We Come Closer to Getting a Sixth Decimal?
(see exhibit 9 above, and exhibit 19 of chapter 10)

A) COMPARISONS -- all entries in units of .000001

		Line A	
x	(*)	Diff	Error
.000	-2	-3	-1
.004	-2	-3	-1
.006	-3	-3	0
.008	-3	-3	0
.014	4	4	0
.016	3	3	0
.020	-1	-2	-1
.026	2	2	0
.028	-1	-1	0
.030	-4	-5	-1
.040	-4	-3	1
.042	2	2	0
.044	-3	-2	1
.046	2	4	2
.050	0	3	3

B) SUMMARY of ERRORS

-4	
-3	
-2	
-1	1111
0	0000000
1	11
2	2
3	3
4	
5	

*****

14H. <u>How far have we come?</u>

Given three points (x,y), we can fit them with

$$y = a + bx + cx^2 .$$

Here cx^2, the new term, is usefully thought of as accounting for bend.

If we take the apparent tilt for each of two pairs of points, the difference of these two tilts is a convenient measure of bend. The difference in tilts for cx^2 will be c times as great as for x^2. By comparing differences in tilt for:

- the data
- for x^2

we can tell what c ought to be. We can equally well use

$$cx^2$$

or

$$c(x-d)^2$$

or

$$c[(x-d)^2 - e]$$

where d or d and e are at our choice as to what is to be subtracted to leave only tilt and level to attend to.

The difference in tilt for x^2 turns out to always be the difference between the unshared x's when the two pairs of points share a point in common.

Given a collection of data points, we can:

- divide them into five portions according to x.
- summarize each portion, often using middle means.
- use the summarizing central points for three of the five portions -- low, middle, and high -- as the basis for fitting bend (and tilt).

To take full advantage of using middle means to summarize portions, we have to go back and forth seeing what modifications to what we have already fitted are needed to polish fit. Fitting either one of bend or tilt may change the apparent best for the other.

*****

14K. Some touches of algebra (optional background)

* why quadratics? *

We have dealt with bends in terms of quadratics in x. Why? It was clear that straight lines have something unique about them -- that they are THE natural way to take out tilt. No such clear feeling about quadratics comes directly to mind.

We could have used other things than quadratics. Sometimes it pays to make other choices.

There is one simple arguemtn, however, in favor of using quadratics as standard bends. It is concerned with how they relate to the straight lines we will use with them to account for tilt and level.

If we expect to deal with U-shaped patterns of data where the bottom of the U -- or the top of the inverted U -- can occur at any x, then the family of standard bends needs to be carefully related to the family of standard tilts that will be used with it. We must be prepared to face U-shaped relations of y to x whose low points fall at different values of x. Yet we want to have very simple standard bends, bends that can be easily described by single numbers. The only way to meet this aim is to use standard tilts to move the low point of the bend back and forth. This means that we depend on sums and differences of standard bends and standard tilts to provide U-shaped partial descriptions with low points wherever we need them.

It would be nice for the subfamily of U-shaped partial descriptions with a given low point to offer the same choices as that for any other low point. This happens for only one choice of a simple family of standard bends, the same choice, fortunately, which most of us would find simplest:

quadratic expressions in x

What happens when we combine a specific quadratic and a specific straight line? A simple example is

$$[(x-10)^2 + 25] + [6x-19] = x^2 - 20x + 125 + 6x - 19$$
$$= x^2 - 14x + 106$$
$$= [(x-7)^2 + 57]$$

which shows that, in this case, we have moved the low point both sideways -- from x = 10 to x = 7 -- and upward -- from y = 25 to y = 57. Except for this rigid displacement, the shape has remained the same. In general, because

$$[(x-a)^2 + c] + [2(a-b)x - a^2+b^2-c+d] = [(x-b)^2 + d],$$

what happens is exactly similar, a rigid displacement, moving the low point from x=a, y=c to x=b, y=d. Clearly we can choose first $2(a-b)$ and then $-a^2+b^2-c+d$ to put the low point wherever we wish.

* difference in tilts for x^2 *

Suppose we are finding one tilt from x_L and x_M and the other from x_M and x_H. If we look at x^2, these tilts are

$$\frac{x_M^2-x_L^2}{x_M-x_L} = x_M+x_L$$

and

$$\frac{x_H^2 - x_M^2}{x_H - x_M} = x_H + x_M$$

and their difference is

$$(x_H + x_M) - (x_M + x_L) = x_M - x_L$$

This is the result we noticed numerically in panel C of exhibit 4.

*****

14P. Problems

1. Assess the bend in the logarithms of the vapor pressure of mercury for temperatures from 0°C to 180°C, using the data given in exhibit 23 of chapter 9. After subtracting the bend, assess the tilt and find the residuals from the partial description combining bend and tilt. Plot these residuals. Discuss. Apply the same partial description to the full range given, finding and plotting residuals. Discuss.

2. Lowering of the freezing point of water by magnesium sulfate. (1,2662) translates as: "For a solution whose concentration gives a normality of 0.01, the freezing-point is lowered 0.02662°C". Data-sets (11): (1,2662), (2,5108), (3,7420), (4,9769), (5,11994), (6,14171), (7,16234), (8,18330), (9,20348), (10,22305), (20,41576). Source: E. H. Loomis 1894. Ueber ein exakteres Vertaliven bei der Bestimmung von Gefrierspunktserniedigungen. Annalen der Physik und Chemie 51: 500-523. Table on page 516. Do a complete job of fitting bend, tilt, and level. Plot residuals. What can you see?

3. Charge of refractive index of fluorite with wave length of infra-red light. (19153,142438) translates as: "For a wave length of **1.9153** microns (19153Å) the observed index of refraction of fluorite was **1.42438**." Data-sets (27): (19153,142438), (19644,142412), (20626,142363), (21608,142317), (22100,142297), (23573,142208), (25537,142092), (26519,142015),

(27502,141969), (29466,141823), (31430,141704), (32413,141608), (35359,141378), (38306,141121), (41250,140850), (44199,140559), (47147,140244), (50092,139902), (53039,139532), (55985,139145), (58932,138721), (64825,137837), (70718,136808), (76612,135672), (82505,134444), (88398,133079), (94291,131612). Source. F. Paschen 1894. Die Dispersion des Fluorits und die Ketteler'sche Theorie der Dispersion. Annalen der Physik und Chemie 51: 812-822. Table on page 822. Problem: Use a desk calculator or a computer to fit a bend, tilt, and level. Find and plot residuals. Summarize your results.

4. Behavior of a new reagent for measuring vitamin A in terms of color production. (94,1020) translates as: "For a concentration of 0.94 micrograms of vitamin A in 5 milliliters of solution the ratio of color to amount of vitamin A, expressed as $E^{1\%}_{1cm}$ at 555 mμ, was 1020." Data-sets (13): (94,1020), (136,1250), (187,1290), (273,1340), (281,1300), (374,1330), (468,1330), (682,1310), (818,1290), (936,1290), (1363,1250), (1871,1180), (2807,1090). (The last translates as: "At 28.07 micrograms in 5 milliliters, the $E^{1\%}_{1cm}$ was 1090.") Source: A. E. Sobel and H. Werbin 1946. Activated glycerol dichlorohydrin. A new colormetric reagent for vitamin A. 18 IECA 570-573. Table II on page 572. Problem: Fit bend, tilt, and level. Find and plot residuals. Summarize results.

5. Given that no $MgSO_4$ makes no reduction in freezing point, check the fitted bend, tilt, and level of problem 2 at zero concentration. State how you could make a fit that would do better at zero concentration.

6. Carry out the better fit just planned. Compare the results, both as to fitted formulas and as to residuals, with problem 2.

7. Electrical resistance of solid bizmuth; increase due to pressure of magnetic field. (0,857) translates as "For a field of 0 grams, the resistance of the smaller bizmuth coil at 18°C was 8.57 ohms." Data-sets (24): (0,857), (5830, (5830,1054), (6310,1074), (6830,1104), (7790,1147), (8880,1206), (10410,1283), (12500,1397), (15710,1560), (20450,1857), (23480,2002), (26820,2150), (27450,2176), (27820,2207), (28370,2237), (29270,2265), (30090,2320), (31270,2377), (32730,2440), (32820,2432), (33300,2478), (35800,2629), (36600,2703), (38900,2856). (The last translates as: "At a field of 38,900 gauss, the resistance was 28.56 ohms".) Source: J. B. Henderson 1894. Einfluss der Magnetisirung und der Temperature auf die electrische Leitfahigkeit des Wismuths. Annalen der Physik und Chemie 53: 912-923. Table 2 on page 919. Problem: Fit bend, tilt and level for resistance against magnetic field. Find and plot residuals. Determine if any simple re-expression would help.

8. Find, from other sources, at least two batches of (x,y) data-sets that interest you and deserve having bend, tilt, and level fitted. Make the fits. Plot the residuals. Summarize the results.

Extra-hard problem

9. Show that a family of standard bends which, when combined with standard straight-line tilts, offers essentially equivalent choices for any fixed low point must be describable by quadratic functions of x.